AIR WAR COLLEGE

AIR UNIVERSITY

DEVELOPING A MORE FLEXIBLE OPERATIONAL TEST AND EVALUATION PROCESS FOR INFORMATION TECHNOLOGY

by

Robert C. Campbell, Jr., Col, USAF

A Research Report Submitted to the Faculty

In Partial Fulfillment of the Graduation Requirements

17 February 2010

Disclaimer

The views expressed in this academic research paper are those of the author and do not reflect the official policy or position of the U.S. government or the Department of Defense. In accordance with Air Force Instruction 51-303, it is not copyrighted, but is the property of the United States government.

Contents

Illustrations

Page

Biography

Colonel Robert C. Campbell, Jr. enlisted in the Air Force in 1980. While on active duty, Colonel Campbell attended Rollins College, Florida, graduating in 1987 with a Bachelor of Science degree in Mathematics/Science. He received his commission through Officer Training School, graduating with honors in 1988. He has served as Commander, 116th Operations Support Squadron, Deputy Commander 116th Operations Group and as a joint action officer at the Air Land Sea Application Center. He has served in several operational and staff positions in the areas of training, tactics, standardization/evaluation and multiservice tactics, techniques and procedures at the squadron, wing and joint levels.

Colonel Campbell graduated from the Air Force Weapons School in 1993, Squadron Officer School in 1994, and the United States Marine Corps Command and Control Systems School in 1996. He received a Master's degree in Education from the University of Oklahoma in 1998 and a Master's degree in Military Operational Arts and Sciences from Air Command and Staff College in 1999.

Colonel Campbell is a senior Air Battle Manager with assignments in strategic air defense, ground theater air control systems and has executed over 1,200 flight hours in the E-3A AWACS and the E-8C Joint STARS aircraft. His awards and decorations include the Defense Meritorious Service Medal with oak leaf cluster, Meritorious Service Medal with two oak leaf clusters, Air Medal, Aerial Achievement Medal with oak leaf cluster, Air Force Commendation Medal with oak leaf cluster and the NATO medal.

Prior to his assignment at Air War College, Colonel Campbell served as Commander of the Air Force Operational Test and Evaluation Center, Detachment 3.

Introduction

Information technology (IT) is ubiquitous in the modern battlespace, existing as standalone programs and embedded in almost every weapons system. Its importance to the warfighter cannot be overstated. That being said, there is a serious disconnect between the current acquisition and testing approach for IT systems and the speed at which they develop. This disconnect is partially due to the Air Force using the same acquisition approach for IT systems as it does for every other major acquisition program. That approach was originally designed for large, relatively static, hardware-based weapons systems. It is characterized by a fairly static, linear process that takes considerable time to complete. Software-based IT systems, however, are unique in that the capabilities they deliver are constantly and rapidly changing during their development and fielding phases.

This disconnect between the rapid, dynamic nature of IT system development and the static, linear standard acquisition approach delays getting vital combat capability into the warfighter's hands. As a result, most IT systems fail to meet expectations. Currently, 25 percent of IT programs fail, 50 percent are delivered late or with less functionality than required and the average business system exceeds budget by almost 100 percent.[1] At the same time, the rapid development of IT systems worldwide offers our potential adversaries the opportunity to catch up, and surpass, our capabilities if the status quo remains. In order to fully capitalize on IT capabilities, the Air Force must revise how it procures and tests these systems.

[1] Steve McConnell, *Professional Software Development: Shorter Schedules, Higher Quality Products, More Successful Projects, Enhanced Careers,*(Boston, MA: 30 June, 2003), xiv.

A key part of the acquisition process is the operational test and evaluation (OT&E) of programs under development. Typically perceived as a final exam that must be passed prior to fielding, OT&E is actually an iterative process executed throughout the acquisition of a given program. As an integral part of the acquisition process, OT&E must also change to reflect the unique nature of IT systems.

This paper will review the problems with the acquisition of IT systems, look at the current "one-size-fits-all" acquisition and testing process and compare it with a new IT acquisition approach recommended by the Defense Science Board (DSB). Finally, this paper will define a more flexible Air Force OT&E process designed to ensure adequate testing without undue delay in fielding these critical IT systems.

Problem Definition

The problem is that "the deliberate process through which weapon systems and information technology are acquired by DoD cannot keep pace with the speed at which new capabilities are being introduced in today's information age."[2] To better understand the unique nature of IT systems, it is important to have a frame of reference. Joint Pub 1-02 defines an information system as "the entire infrastructure, organization, personnel, and components for the collection, processing, storage, transmission, display, dissemination, and disposition of information."[3] For the purpose of this paper, IT is defined as any system of hardware and/or software whose primary purpose is the manipulation of information. Based on that definition, the battlespace is replete with IT systems and destined to become even more so. The problem can be better

[2] Office of the Under Secretary of Defense for Acquisition, Technology and Logistics, Defense Science Board, *Department of Defense Policies and Procedures for the Acquisition of Information Technology*. (Washington DC: March 2009), 1.
[3] Office of the Secretary of Defense, *JP 1-02, DOD Dictionary of Military and Associated Terms.* (Washington, DC: 19 August 2009), 263.

understood by looking at how large these IT systems have grown, how much they cost, and how much time it takes to field them.

The disconnect becomes exponentially more critical as IT systems continue to proliferate across the battlespace. "Whereas in 1970 software accounted for about 20 percent of weapon system functionality, by 2000 it accounted for as much as 80 percent and today can deliver 90 percent or more of a system's functionality."[4] Additionally, even as the functionality of software grows, the complexity of that software is growing even faster. For example, the number of lines of code found in Microsoft Windows has grown ten-fold from Windows 3.1 in the early 90s, to over 50 million for Windows Vista in 2007.[5] This growth is mirrored in today's weapons systems. "The executable source lines of code (ESLOC) within weapons systems, such as missiles, ships, and aircraft have grown from a few thousand to tens of millions."[6]

The problem of expansion is mirrored in the cost of these systems as well. In 1990, DoD recognized that the cost of acquisition of IT systems was ballooning out of control. DSB reports found only 16% of all IT systems were on budget and on time while 53% were late and over budget, typically by over 89%.[7] That problem has not been resolved. A 2008 GAO review "concluded that 48 percent of the federal government's major IT projects have been re-

[4] Office of the Under Secretary of Defense for Acquisition, Technology and Logistics, Defense Science Board, *Department of Defense Policies and Procedures for the Acquisition of Information Technology*. 6.
[5] James Larus, "Spending Moore's Dividend," *Communications of the Association for Computing Machinery* 52, no. 5, (May 2009): 64.
[6] Office of the Under Secretary of Defense for Acquisition, Technology and Logistics, Defense Science Board, *Department of Defense Policies and Procedures for the Acquisition of Information Technology*. 14.
[7] Office of the Under Secretary of Defense for Acquisition and Technology, Defense Science Board, *Task Force on Defense Software*. (Washington DC: November 2000), 11.

baselined…[and] 51 percent were re-baselined at least twice."[8] These re-baselinings increased costs.

Case in point is the Joint Tactical Radio System (JTRS), a series of modular radios designed for command posts, ground vehicles and a range of aircraft. Originally estimated to cost $3.5 billion and to begin fielding in 2001, this program has ballooned to over $6 billion dollars and is almost ten years late. Additionally, these problems forced the services to spend an additional $6.1 billion on legacy radios.[9] The issue is not restricted to just cost growth. IT systems are also hampered by fielding delays.

IT systems as a whole are suffering with delays in fielding while at the same time the pace of software capability growth is increasing. On average, most DoD IT systems are almost two years behind schedule in delivering initial operational capability and 12 percent are over five years late.[10] The result is often the warfighter is given increasingly obsolete systems on an increasingly delayed timeline. Couple this with the fact that our adversaries can, and do, take advantage of modern IT systems themselves while we burden ourselves with out-of-date software, and the problem becomes even more acute.

IT systems are increasing in number, increasing in complexity, increasing in cost and are increasingly delayed in fielding. While some of the blame for these problems is attributed to technological difficulties, the primary reason can be assigned to the use of the standard,

---

[8] Office of the Under Secretary of Defense for Acquisition, Technology and Logistics, Defense Science Board, *Department of Defense Policies and Procedures for the Acquisition of Information Technology*. 44.

[9] Bob Brewin, "Defense Radio Project Not Practical or Affordable, GAO Says," *Nextgov Technology and the Business of Government*, (18 August 2008): 1, http://www.nextgov.com/nextgov/ng_20080818_4317.php?oref=search.

[10] US Government Accountability Office, *GAO-08-782, Better Weapon Program Outcomes Require Discipline, Accountability, and Fundamental Changes in the Acquisition Environment*. (Washington DC: June 3 2008), 5.

cumbersome acquisition process. This process is too inflexible for the dynamic nature of IT systems and the result is an increase in costs and fielding delays. To better understand this issue, first we will examine the standard acquisition process.

Current Acquisition Model

A weapons system designed for one use can often be adapted to serve another purpose. Case in point is the use of modern targeting pods on strike aircraft to perform non-traditional intelligence, surveillance and reconnaissance. If a legacy system cannot be adapted to fill a defined capability gap, or changes in doctrine or training cannot mitigate that gap, then a new weapons system must be developed. This, however, is not a simple task. Modern weapons systems are increasingly complex and expensive. The Defense Acquisition System was created to manage the development of these complex and expensive acquisition programs.

The Defense Acquisition System is designed to acquire mission capability based on requirements identified through the Joint Capabilities Integration and Development System (JCIDS). JCIDS attempts to link missions defined in the national military strategy to specific weapons systems. Combatant commanders identify capabilities required to accomplish their mission, and these capabilities are prioritized at the Joint Requirements Oversight Council (JROC). The objective is to ensure the services acquire specific capabilities designed to meet the joint warfighter's requirements.[11]

Once the JROC approves those specific requirements, and a new material solution is required, the services use the Defense Acquisition System to procure weapons systems meeting those specified requirements. The Defense Acquisition System model is illustrated in Figure 1.

[11] Department of Defense, *CJCSM 3170.01C, Operation of the Joint Capabilities Integration and Development System.* (Washington, D.C.: March 2009), A-1 – A-4.

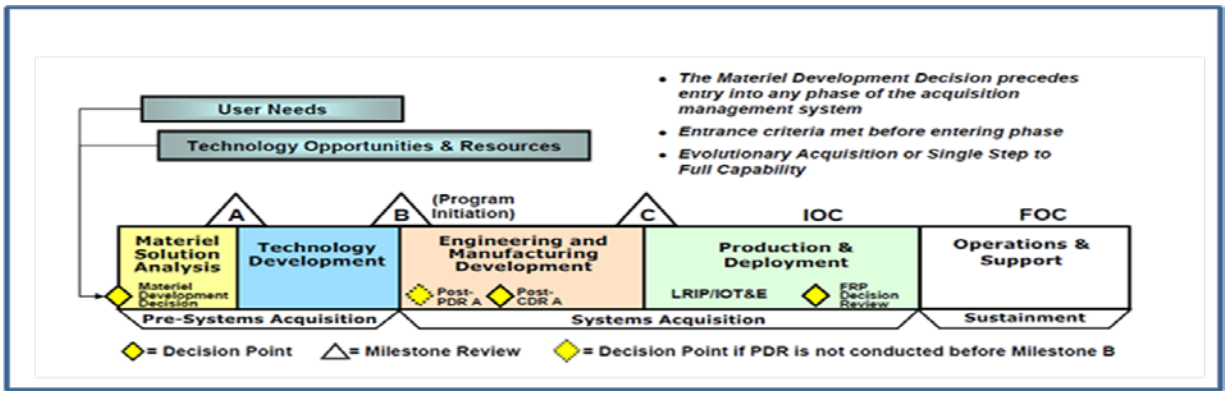

Figure 1. Current Acquisition Model (Reprinted from DODI 5000.02, *Operation of the Defense Acquisition System*, December 2008.)

Each stage has a specific purpose and is supported by a specific milestone. The Milestone A decision reviews potential solutions identified in the Material Solution Analysis phase as defined by the Initial Capabilities Document (ICD). The Technology Development phase moves the nascent weapons system from concept to prototype and ends with a Milestone B decision that transitions the program into development. The Engineering and Manufacturing Development phase spells out specific capabilities in the Capabilities Development Document (CDD) that defines the key performance parameters the system must meet to be approved for production. This phase ends with a Milestone C decision, supported by a Capabilities Production Document (CPD). After Milestone C, a program usually declares initial operational capability (IOC) and begins to transition to sustainment.[12] This process, from requirements definition to initial operational test and evaluation (IOT&E) takes years to complete.

Current Operational Test and Evaluation Approach

Supporting this acquisition model is a series of operational tests and evaluations that determine the operational effectiveness and suitability of the program under development. For

[12] Department of Defense, *DODI 5000.02, Operation of the Defense Acquisition System*. (Washington, D.C.: December 2008), 14-22.

Air Force-led acquisition programs, the Air Force Operational Test and Evaluation Center (AFOTEC) accomplishes test planning, design and execution to provide information to support acquisition decisions. OT&E is defined as "a field test, under realistic combat conditions…of weapons, equipment, or munitions for the purpose of determining the effectiveness and suitability…for use in combat by typical military users."[13] These operational tests range from simple Operational Assessments to full-blown Initial Operational Test and Evaluations (IOT&E).

The primary OT event conducted by AFOTEC is an IOT&E. This test is conducted with operational warfighters in as realistic an operational environment as possible and with a production-representative system. The intent is to estimate the system's overall operational capability as determined by its effectiveness, suitability and mission capability. An IOT&E also analyzes the organizational, training, doctrine and tactics requirements of the system to ensure the warfighter receives the complete, sustainable capabilities originally described in the CPD.[14]

The level of effort and time involved in planning and executing these tests represents a significant allocation of resources. The standard test timeline from start to finish for most AFOTEC OT&E events is notionally 18-months. AFOTEC defines a rapid test approach as beginning major test activities within 12-months of program inception.[15]

This timeline represents a standard approach for standard acquisition programs. Due to the dynamic nature of IT systems, DoD has recognized this standard approach isn't sufficient and tasked the DSB to review the acquisition of IT systems and identify improvements. The DSB issued its report in March 2009. Of the four major recommendations, the most significant was to

---

[13] Department of the Air Force, *AFI 99-103, Capabilities-Based Test and Evaluation.* (Washington, D.C.: 26 February 2008), 73.
[14] Air Force Operational Test and Evaluation Center, *AFOTECI 99-103, Conduct of Operational Test and Evaluation.* 6th Edition. (Kirtland AFB, NM: 12 February 2009), B-14 – B-15.
[15] Ibid, 1-20.

develop an IT-unique acquisition model more responsive to the dynamic challenges of the IT arena.[16] That approach represents the latest in several attempts to streamline the acquisition and testing of IT systems. Two of these attempts are discussed below.

In 1996, Congress attempted with the Clinger-Cohen Act (CCA) to mandate an over-arching management perspective on the acquisition of IT systems. CCA centralized procurement of IT systems by creating a Chief Information Officer responsible for confirming all CCA requirements were met before acquiring an IT system. The goal was to centrally manage IT acquisitions, share common systems and move IT acquisitions from an agency-dependant process to a centralized approach.[17] It hasn't worked as planned. Instead, CCA has basically levied additional administrative requirements without improving the acquisition process. The first specific action to do so came from within the operational test community.

In 2003, the Director, Operational Test and Evaluation (DOT&E) issued a directive entitled "Guidelines for Conducting Operational Test and Evaluation for Software-Intensive System Increments". These guidelines recognized the evolutionary acquisition model for developing and fielding increments of capabilities and tied them to a risk assessment of each increment. If an increment represented a low or moderate level of risk to the system as a whole, then the level of operational testing could be lowered correspondingly.[18] While this guidance represented a

[16] Office of the Under Secretary of Defense for Acquisition, Technology and Logistics, Defense Science Board, *Department of Defense Policies and Procedures for the Acquisition of Information Technology*. 61-68.
[17] "Information Technology Management Reform Act of 1996" Title 40, *U.S. Code, §139, et. seq., Clinger-Cohen Act*. (Washington, D.C.: 10 February 1996).
[18] Office of the Director of Operational Test and Evaluation, *Guidelines for Conducting Operational Test and Evaluation for Software-intensive Systems*. (Washington, D.C.: 16 June 2003), 1-2.

change to the operational test approach and timeline, it did not go far enough to address the overall acquisition timeline. Further improvements were necessary.

Information Technology-Unique Acquisition Model

In 2008, Congress directed the DoD to review the acquisition of IT systems. In March, 2009, the DSB released a report that contained four basic recommendations. First, strengthen the roles and responsibilities of the DoD Chief Information Officer in regard to the acquisition of IT systems. Second, consolidate all acquisition oversight of IT systems to the DoD CIO. Third, improve the subject matter expertise of IT acquisition officers by hiring more experienced and trained personnel. The most important recommendation; however, was to implement a new IT-unique acquisition model to replace the standard DODI 5000.2 model. This recommendation would change the approach used to acquire IT systems with the goal of developing and fielding IT increments within 18 months. That model is illustrated below.[19]

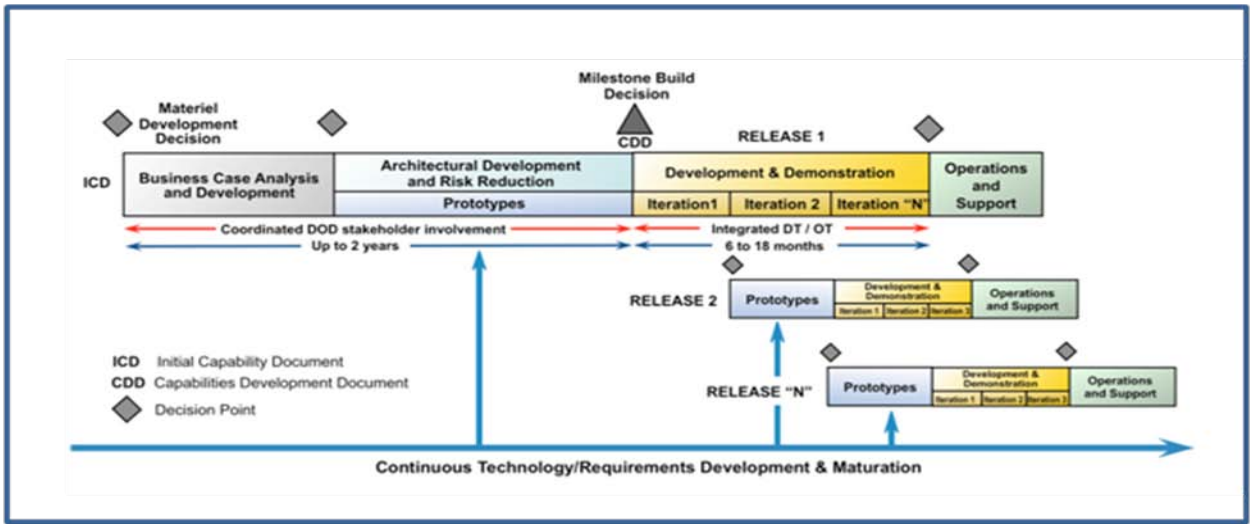

Figure 2. Defense Science Board Recommended Acquisition Model for IT Systems (Reprinted from *Department of Defense Policies and Procedures for the Acquisition of Information Technology*, March 2009.)

[19] Office of the Under Secretary of Defense for Acquisition, Technology and Logistics, Defense Science Board, *Department of Defense Policies and Procedures for the Acquisition of Information Technology*. 61-68.

This IT-unique acquisition model adapts the incremental approach defined in the latest DODI 5000.2 with simplified program milestones. Requirement documents are still required early in the process, but those requirements are expected to evolve so that "'desired capabilities' can be traded-off against cost and initial operational capability to deliver the best capability to the field in a timely manner. A modular, open-systems methodology is required, with heavy emphasis on 'design for change', in order to rapidly adapt."[20] Essentially, this approach allows for shifting specified capabilities from one increment to another to quickly field capabilities that have increased in priority or demonstrated maturity sooner. It gives more flexibility to the program manager to decide, with input from the warfighter, which capabilities to deliver based on need and performance. This flexibility is inherent in the process and does not require constant revisits to the JROC for requirement changes. The Senate, in the FY2010 Defense Authorization Act, directed DoD to "develop and implement a new acquisition process for information technology systems…based on the recommendations…of the March 2009 report of the DSB…. [This approach would] be designed [for]…multiple, rapidly executed increments or releases of capability to support an evolutionary approach."[21]

This directive, while an excellent start, really only addresses the acquisition portion of IT systems. The DSB report makes little mention of how to test within this new approach. It does acknowledge the critical role testing plays. "Testing methodologies and procedures need to be engaged early and often in the acquisition process, with integrated and continuous development

[20] Office of the Under Secretary of Defense for Acquisition, Technology and Logistics, Defense Science Board, *Department of Defense Policies and Procedures for the Acquisition of Information Technology*. xi.

[21] "National Defense Authorization Act for Fiscal Year 2010" *§804, et. seq.*,(Washington, D.C.: 28 October 2009), Section 804.

and operational testing practiced during the development and demonstration phase for each capability release."[22] The proposal, however, does not describe how to accomplish this integration. The next step is to change the operational test approach for IT systems.

New Capabilities-Defined Test Model for IT Systems

Congress determined a new approach was needed for the timely acquisition of IT systems. The Air Force should follow up this decision by defining a new approach for the OT&E of those systems. This new approach is not ground-breaking in concept. In fact, the elements already exist. This new approach merely integrates them and codifies them into a more responsive OT approach. This new approach consists basically of two main elements; first, a risk-assessment test format based on the DOT&E guidelines for software-intensive systems and second, a more responsive reporting approach for OT reports. Some of these elements have been used, or are being used at AFOTEC, although only on an ad hoc basis. This proposal codifies them into a comprehensive, formal approach for testing IT systems that is as flexible as the new acquisition approach.

The DSB recognized that using the standard acquisition model for IT isn't effective. Their recommendation was an incremental acquisition model that allows capabilities to shift amongst increments as the priority or readiness of those capabilities matures. The model's intent is to get relevant, tested capabilities into the warfighter's hands as quickly as possible. AFOTEC shares the same basic intent. Their mission to "deliver warfighting capabilities faster and with more

[22] Office of the Under Secretary of Defense for Acquisition, Technology and Logistics, Defense Science Board, *Department of Defense Policies and Procedures for the Acquisition of Information Technology*. xi.

confidence"[23] (AFOTEC/CC brief) is conditioned on their ability to plan, execute and report OT&E as efficiently as possible. Unfortunately, just as the acquisition community has been handcuffed by using the standard acquisition model for IT systems, AFOTEC has also been hampered by using their standard OT&E process for the more dynamic IT programs. This standard process is very linear and depends on a formalized test and evaluation strategy that is defined early in the development of an IT program and is characterized by discrete test plans and reports for each test event. This process is illustrated below.

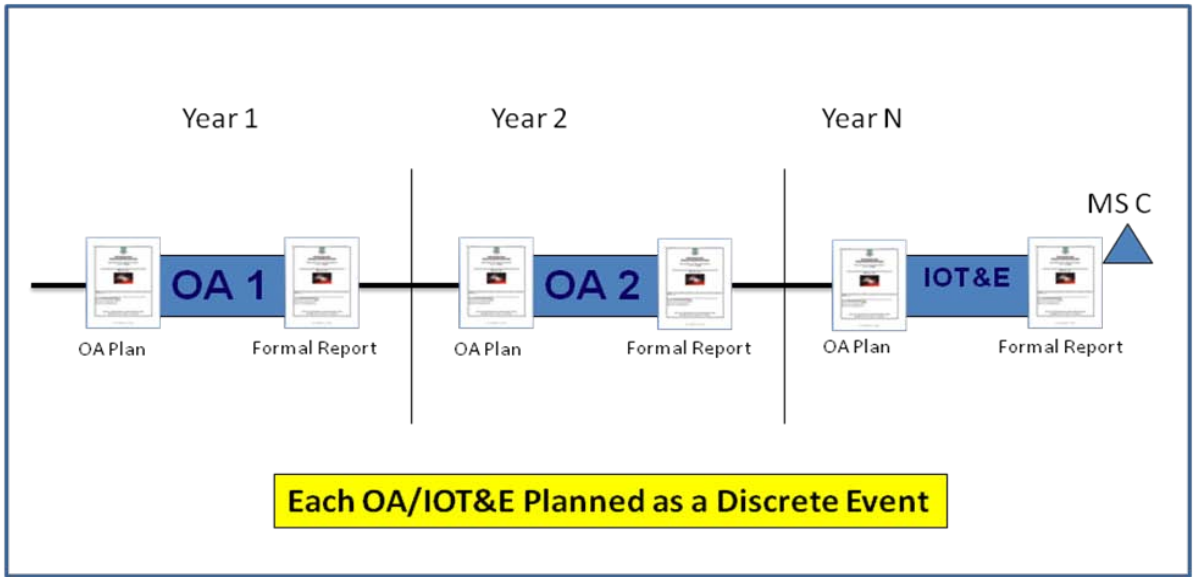

Figure 3. Current Discrete OT Planning Model (Adapted from briefing, Rich Brunson, AFOTEC, Det 3, *Test Driven Development*, 15 May 2009.)

The primary problem with this approach is that substance becomes the victim of style. Each OT plan is a formalized event that can take months to complete and gain approval. Additionally, each report is just as formalized and can sometimes take as much time to write and staff for approval as it took to conduct the operational test event itself. Fundamentally, these are

[23] Major General Steve Sargeant, "AFOTEC Initiatives to Improve Operational T&E" (unclassified Air Force Operational Test and Evaluation brief, Manhattan Beach, CA, 14 October 2009).

administrative issues that can and should be reviewed. The major issue, however, is that as this formalized process is taking place, the IT program under evaluation is changing from increment to increment. Each change made to move a capability from one increment to another requires a corresponding change in the test plan. Since the test plan process is already lagging the development phase, this places the test plan even more behind reality. As a result, while the formalized IOT&E event is planned and coordinated, increments of capability are finalized and fielded without dedicated OT&E. Often, by the time of the IOT&E, up to 90% of an IT program can be in the field being used without an operational test to validate the warfighter is getting what he asked for.[24] The test structure itself needs to be more flexible. AFOTEC has recognized this and is working with the acquisition community to develop new test models.

As the acquisition model has changed to allow more flexible shifting of capabilities from increment to increment, the test structure needs to change to address those shifts. The first element of this more flexible OT&E construct should be to replace the discrete increment by increment stand-alone OT process with an open-ended integrated DT/OT approach that executes continuously as capabilities are developed and ready for fielding. Guided by one over-arching OT plan that covers all capabilities, regardless of which increment they are delivered in, this approach retains the flexibility necessary to shift as the program itself changes. This approach is illustrated in Figure 4.

[24] Ibid.

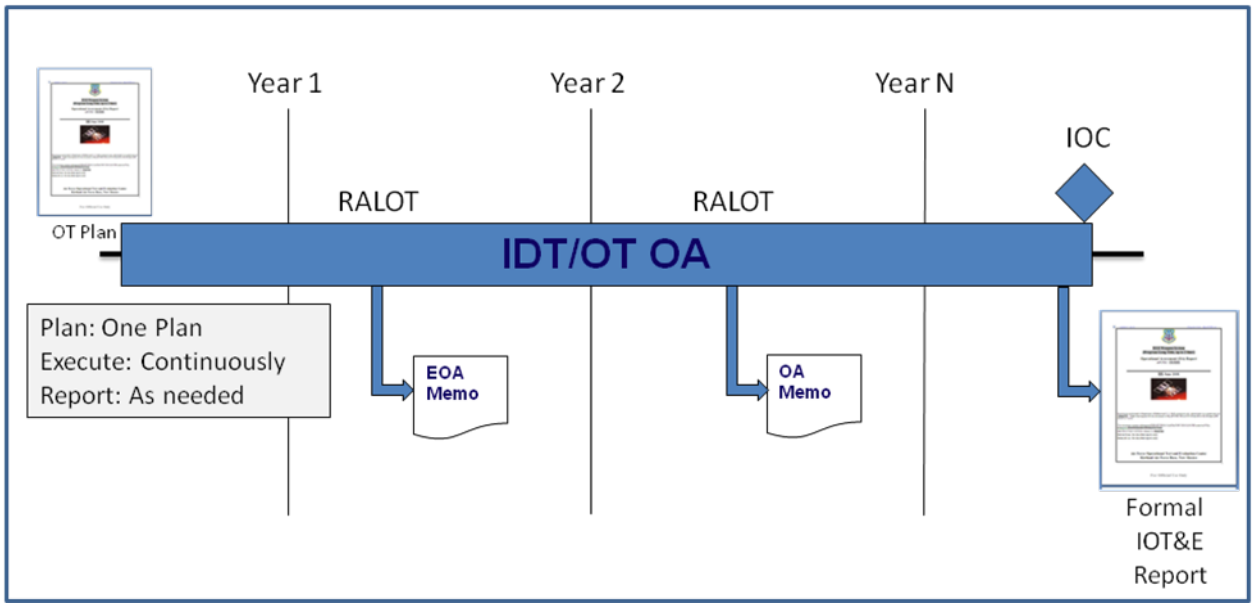

Figure 4. Proposed Capabilities-Defined OT Model (Adapted from briefing, Rich Brunson, AFOTEC, Det 3, *Test Driven Development*, 15 May 2009.)

The fundamental change is one of structure and philosophy. Instead of a predetermined test defined by a predetermined number and type of tested capabilities, this approach involves numerous test events of different types defined in time and structure by the type of capabilities developed. Based on the new acquisition model, these capabilities can and will shift from increment to increment. This capabilities-defined test approach will also shift as the delivered capabilities change. This type of test approach was under development by AFOTEC Detachment 3's Technical Advisor as recently as July, 2009.[25]

The second aspect of this change in structure is the type of test to conduct for a given increment. Instead of a predetermined type of test, the test event itself will also be defined by the nature of the delivered capabilities. Using the 2003 DOT&E Guidelines as a foundation, this

[25] Rich Brunson (Technical Advisor, AFOTEC Detachment 3), telephonic interview by the author, 14 October 2009.

approach uses a risk-assessment of the delivered capability to determine the level of test to be performed.

As mentioned earlier, in 2003, DOT&E recognized the growing acceptance of incremental development of IT systems and developed OT&E guidelines for conducting focused test and evaluation based on the level of risk that increment introduces to the overall enterprise. The concept is that if a new increment delivered capabilities with relatively low risk to the system, then the subsequent level of OT should also be low risk. The objective was "to provide a method for determining levels of operational testing appropriate to the risk posed by specific system increments."[26] The process consists of four steps: prepare risk assessment; determine appropriate level of OT&E; develop an OT&E plan for that level of test; then execute and report test results. The core of this approach lay in the risk assessment. This assessment was composed of two parts: an analysis of the likelihood of success of an increment and an understanding of mission impact of increment failure; and a definition of the amount of OT&E that provides sufficient assurance that the risk will be mitigated.[27]

To implement the DOT&E guidelines, AFOTEC developed a risk assessment, level of test tool (RALOTT) which applied a systematic approach for determining the risk of new increments and recommended an appropriate level of test for each increment. However, if the RALOTT recommended anything other than a level 1 test, the entire test event would have to undergo the normal AFOTEC test planning process and a dedicated OT report would have to be written and staffed. Again, this process has not proven to be responsive enough for IT systems.

---

[26] Office of the Director of Operational Test and Evaluation, *Guidelines for Conducting Operational Test and Evaluation for Software-intensive Systems*. (Washington, D.C.: 16 June 2003), 2.
[27] Ibid., 2.

Another difficulty with the approach of the DOT&E guidelines is that they apply only to increments of a program and not to the "core increment" which still required full operational testing.[28] Unfortunately, the core increment is usually fielded first and would require a full-blown AFOTEC OT event, which negates the timeliness of the process. What is needed is a more responsive approach.

AFOTEC should coordinate with DOT&E to update these guidelines to remove the restrictions on the "core increment" and replace it with a risk assessment level of test for each increment as it is delivered, followed by a formalized IOT&E for the entire system once the final increment is produced. This would allow continuous OT throughout the development of the program which could identify problems as early as possible and allow corrective actions at a relatively inexpensive point. This approach would provide actionable information to the decision-makers on the effectiveness, suitability and mission capability of each increment as it is being developed.

The other services, and even industry, have embraced this concept of risk-based testing. The Army, as lead operational test agency for the Net-Enabled Command Capability (NECC), a joint command and control software suite, describes "Risk Analysis/Level of Test (RALOT)" as a means to determine what level of operational test is required for each capability module delivered. The USMC, as a signatory to the NECC Test and Evaluation Master Plan, has also tacitly accepted the RALOT concept.[29] The Navy includes the RALOT concept in its operational test guide and even expands its use beyond just IT systems. However, it restricts its application

[28] Ibid., 1.
[29] Joint Program Management Office, *Net-Enabled Command Capability Test and Evaluation Master Plan*. (Arlington, VA: 28 January 2009), 2.2.3.4.

to modifications of post-IOT&E legacy systems, not to include emerging programs.[30] Even Microsoft has endorsed the principle of risk analysis in deciding whether or not to fix bugs discovered during their tests of new software. Explaining why software often ships with known bugs, Alan Page, one of the authors of *How We Test Software at Microsoft*, states that "many bugs aren't worth fixing…sometimes, the risk and investment just aren't worth it."[31] Page describes the issues of severity, frequency and impact as the factors that determine whether or not a known bug is fixed based on a risk analysis of the software under test.[32] Since modern IT systems are usually delivered in increments, the concept of risk analysis can easily be applied to each increment with regards to the risk that increment presents to the entire system. This is the first step of the capabilities-defined OT approach: a new open-ended test structure based on risk assessment of each increment. The second step is a more responsive reporting approach.

As with the current structure of OT events, the current reporting of OT results is overly formalized and lags fielding decisions. What is needed is timely, actionable information the decision-maker can use when deciding whether or not to field a capability. That information currently is contained in AFOTEC's formalized OT report. Unfortunately, that report takes a considerable amount of time to write, staff, brief and release. Often, fielding decisions are made prior to the report completion. One way to resolve this dilemma is to define exactly what information the decision-maker needs to make his decision, and deliver it quicker. One approach is to use a product already produced by AFOTEC.

---

[30] Commander Operational Test and Evaluation Force, *COMOPTEVFORINST 3980.1, Operational Test Director's Manual*, 6-55.
[31] Alan Page, "Why Bugs Don't Get Fixed", *Notes and Rants*, http://blogs.msdn.com/alanpa/archive/2009/09/30/why-bugs-don-t-get-fixed.aspxhy bugs don't get fixed.
[32] Ibid.

The Evaluation Summary Chart (ESC) is a matrix developed during initial test planning that lays out the Critical Operational Issues (COIs) and specific operational capabilities in order to define Measures of Effectiveness (MOE) and Measures of Suitability (MOS). In short, the ESC provides a vehicle for demonstrating program performance.[33] A notional ESC is illustrated below.

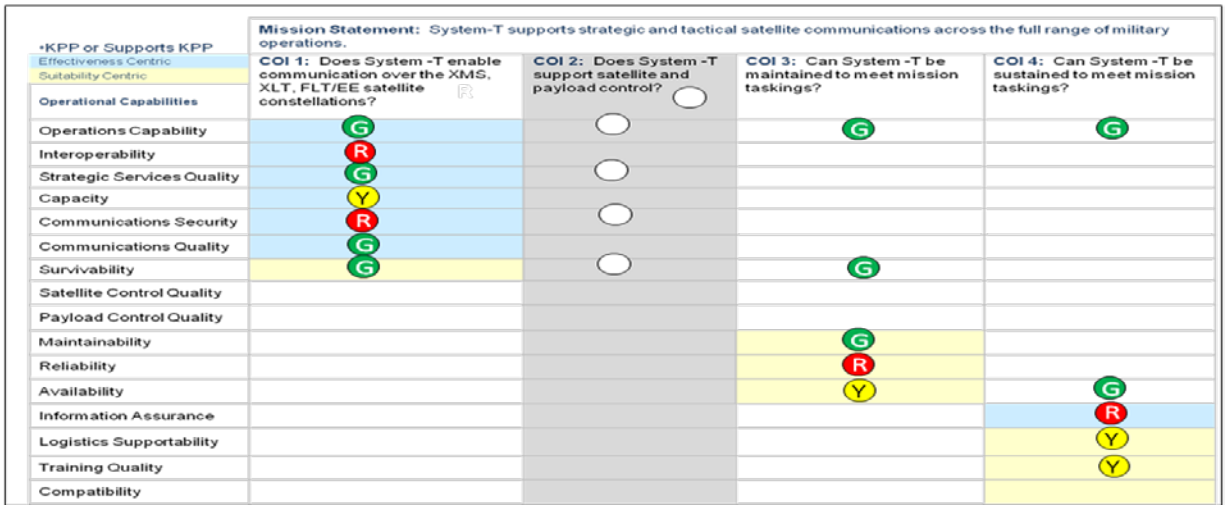

Figure 5. Notional Evaluation Summary Chart (Reprinted from briefing, Rich Brunson, AFOTEC, Det 3, *Test Driven Development*, 15 May 2009.)

The ESC provides an operationally relevant, decision quality snapshot of performance measured against the requirements defined in the capabilities documents. It is operationally relevant because it directly relates warfighter requirements to capabilities inherent in the design of the program. It is decision quality because it easily conveys to the decision maker what the warfighter can and cannot do with the system under test. The ESC provides the decision maker with insight for making engineering tradeoffs and allocating resources to fix shortcomings in system capabilities. Unfortunately, that snapshot is just that; performance as measured at a single

[33] Air Force Operational Test and Evaluation Center, *AFOTECI 99-103, Conduct of Operational Test and Evaluation*, 2-12 – 2-13.

point of time. This vehicle can also be used to provide continuous updates on program status as it evolves throughout the flexible capabilities-based testing described previously. As a program proceeds from one OT event to another, this ESC would be updated and placed on a secure website for both the program manager and milestone decision authority to review. As integrated developmental test/operational test (IDT/OT) events occur in between formal OT events, the ESC can be updated almost immediately by the test director, pending approval by AFOTEC/A3, providing critical data independent of the formal AFOTEC OT report.

The formal report would still be produced and disseminated to record overall program performance, but it would be supplemented with the instant ESC updates and less formal interim status reports and memos as indicated by the risk assessment determination done in step one. Coupling this more flexible reporting approach with the dynamic risk-assessment format allows AFOTEC to perform its critical operational test and evaluation function on rapidly changing IT systems, and get the information out to the decision-makers before these systems are approved for fielding.

While this approach is more flexible and responsive than the standard test construct, it is not without potential risks. Legitimate concerns about less stringent tests, false information in loosely controlled reports and never-ending involvement have been raised and must be addressed. AFOTEC has built a hard-earned reputation for conducting well-constructed, tightly-executed and evenly-reported operational tests. In some AFOTEC staffer's eyes, switching to the proposed capabilities-defined test approach potentially challenges the essence of what AFOTEC stands for. Those concerns can be mitigated by applying the same systematic oversight to the new testing approach as was applied for the standard process. For instance, concerns about a less-stringent test construct can be resolved by ensuring members of AFOTEC/A3 and A2/A9

are actively engaged in the RALOT process itself. Additionally, concerns about the quality of the information released can be eased by ensuring the executing detachment's Technical Advisor, AFOTEC/A3 and even A2/A9 if necessary, reviews any interim report or ESC update before dissemination. Finally, the concern of being trapped into a never-ending involvement with increment after increment of a new system can be alleviated by working with the program manager to define the exact point when the program transitions from acquisition to sustainment. All of these concepts have been utilized in the past. Integrating them into a codified test approach offers a more flexible response to the dynamic IT world.

A version of this approach has been done at AFOTEC in the past. Over the past five years, AFOTEC Detachment 3 planned, executed and reported on the operational test and evaluation of the Integrated Strategic Planning and Analysis Network (ISPAN), a network-centric mission planning and execution system designed for US Strategic Command. Over the course of three and a half years, AFOTEC informed nine fielding decisions based on integrated DT/OT activities; accomplished risk assessments for nine test events; operationally tested capability that was fielded every six months; reported many of those events via a less formal Assessment Memorandum within 10 days of last test event; and finalized the program with a comprehensive IOT&E and formal test report.[34] This approach has proven flexible enough for a dynamic IT system like ISPAN, although it has not been fully accepted at AFOTEC. The IT systems of the future are destined to be even more dynamic than ISPAN. To enhance responsiveness, the Air Force should standardize this approach now and couple it with ESC-based informal reporting maintained on-line to provide continuous feedback on program performance prior to fielding.

[34] Brunson interview

Conclusion

The current acquisition and testing system is optimized for large scale, hardware-based weapons systems. Capabilities delivered by software-intensive IT systems, however, are much more dynamic than these hardware systems and the current acquisition process slows down fielding these capabilities. OT&E is an integral part of the acquisition process. Within the acquisition system itself, the current Air Force OT&E process also is optimized for hardware-based weapons systems. Unless a change is made in the OT&E process, it will be increasingly difficult for the Air Force to field critical IT capabilities in a timely manner.

The DSB recognized the disconnect between using a static acquisition process for acquiring dynamically changing IT systems. It has recommended a unique, incremental acquisition model for IT capabilities. That new acquisition model requires a correspondingly dynamic and flexible OT&E approach to ensure these critical systems have been adequately tested on a more responsive timeline. That approach should be an open-ended OT construct determined in structure and schedule by the capabilities being delivered in a given increment. Since those capabilities are constantly shifting, the OT approach should be flexible enough to shift as well. The approach should combine a risk-assessment model to determine the appropriate level of OT events along with a more flexible reporting process of on-line performance reports and less formal interim summary reports and memos. AFOTEC needs to embrace this more dynamic proposal to retain the timeliness and relevance of its critical OT reporting. Changing the acquisition process was an important first step. Changing the OT&E approach is an equally important second step.

Bibliography

Air Force Instruction (AFI) 99-103. *Capabilities-Based Test and Evaluation*, 2008.

Air Force Operational Test and Evaluation Center Instruction (AFOTECI) 99-103. *Conduct of Operational Test and Evaluation*, 2009.

Brewin, Bob. "Defense Radio Project Not Practical or Affordable, GAO Says." *Nextgov Technology and the Business of Government*, August 2008. http://www.nextgov.com/nextgov/ng_20080818_4317.php?oref=search. (accessed 17 November 2009).

Brunson, Richard. "Test Driven Development." Unclassified Air Force Operational Test and Evaluation, Detachment 3 brief, draft, 15 May 2009.

Chairman, Joint Chiefs of Staff Memorandum (CJCSM) 3170.01C. *Operation of the Joint Capabilities Integration and Development System*, Washington, D.C., 2007.

Chairman, Joint Chiefs of Staff Joint Publication *JP 1-02. DOD Dictionary of Military and Associated Terms*, August 2009.

Commander, Operational Test and Evaluation Force Instruction (COMOPTEVFORINST) 3980.1. *Operational Test Director's Manual*, 2008.

Department of Defense (DOD) Instruction 5000.02, *Operation of the Defense Acquisition System*, Washington, D.C., 2008.

Director of Operational Test and Evaluation (DOT&E) *Guidelines for Conducting Operational Test and Evaluation for Software-intensive Systems*, Washington, D.C., 2003.

Information Technology Management Reform Act of 1996. US Code. Title 40, §139, et. seq., Clinger-Cohen Act, 1996.

Joint Program Management Office. *Net-Enabled Command Capability Test and Evaluation Master Plan*, Arlington, VA, 2009.

Larus, James. "Spending Moore's Dividend." *Communications of the Association for Computing Machinery* 52, no. 5, May 2009.

McConnell, Steve, *Professional Software Development: Shorter Schedules, Higher Quality Products, More Successful Projects, Enhanced Careers*, Boston, MA: Addison-Wesley Professional, 2003.

Office of the Under Secretary of Defense for Acquisition, Technology and Logistics. Defense Science Board, *Department of Defense Policies and Procedures for the Acquisition of Information Technology*, Washington D.C., 2009.

Office of the Under Secretary of Defense for Acquisition and Technology. Defense Science Board, *Task Force on Defense Software*, Washington D.C. November 2000.

Page, Alan. "Why Bugs Don't Get Fixed." *Notes and Rants*, September 2009. http://blogs.msdn.com/alanpa/archive/2009/09/30/why-bugs-don-t-get-fixed.aspxhy bugs don't get fixed (accessed 18 January 2010).

Sargeant, Major General Steve. "AFOTEC Initiatives to Improve Operational T&E." Unclassified Air Force Operational Test and Evaluation brief, Manhattan Beach, CA, October 2009.

United States Congress. “National Defense Authorization Act for Fiscal Year 2010” *§804, et. seq.*,Washington, D.C. October 2009.
United States Government Accountability Office. *GAO-08-782, Better Weapon Program Outcomes Require Discipline, Accountability, and Fundamental Changes in the Acquisition Environment*, Washington D.C. June 2008.